

— Supplementary Material —  
Comparison of Subjective Methods for Quality  
Assessment of 3D Meshes in Virtual Reality

Yana Nehmé      Jean-Philippe Farrugia      Florent Dupont  
Patrick Le Callet      Guillaume Lavoué

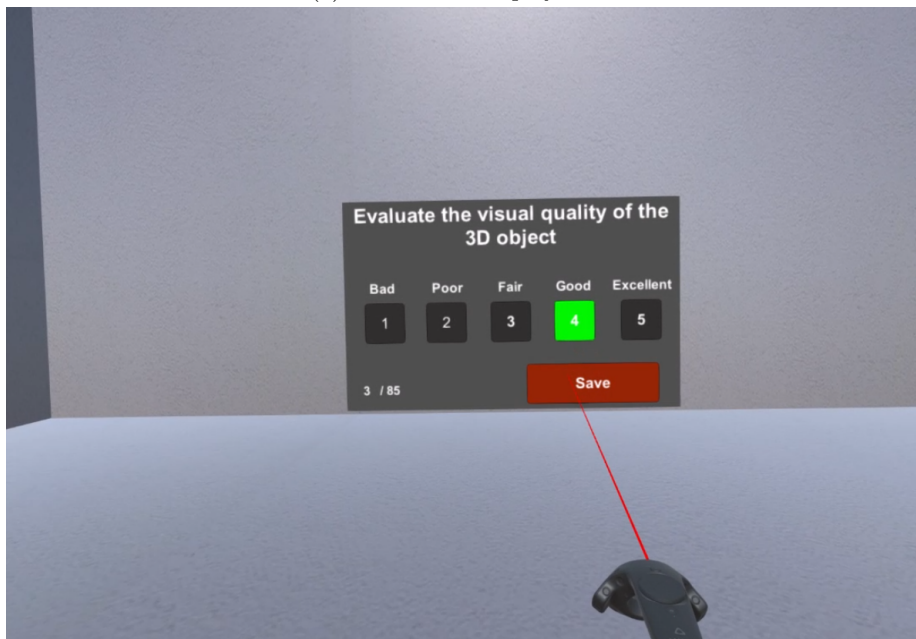
This supplementary material is organized as follows. Section 1 shows snapshots of the experimental environment for the ACR-HR, DSIS and SAMVIQ tests. Section 2 illustrates the DMOS and MOS for the two groups of subjects involved in the ACR-HR and DSIS sessions and those for the SAMVIQ test participants. We also provide the confidence intervals of the computed DMOS/MOS. Section 3 describes an intra-rater reliability analysis among the ACR-HR and DSIS tests for the two groups of subjects. Section 4.1 compares confidence intervals (CI) from the ACR-HR and DSIS methodologies between the 2 groups of subjects, while section 4.2 assesses the evolution of the confidence intervals according to the number of subjects, for the 3 tested methods. Finally, we provide, in section 5, a summary of all the experimental details of the three tested methods.

## 1 Subjective Experiment

In designing our subjective experiment, we opted to ensure a user experience and quality of experience (QoE) in fully immersive virtual environment (VE). Figure 1, Figure 2 and Figure 3 show snapshots of the ACR-HR, DSIS and SAMVIQ session respectively.



(a) The stimuli display room

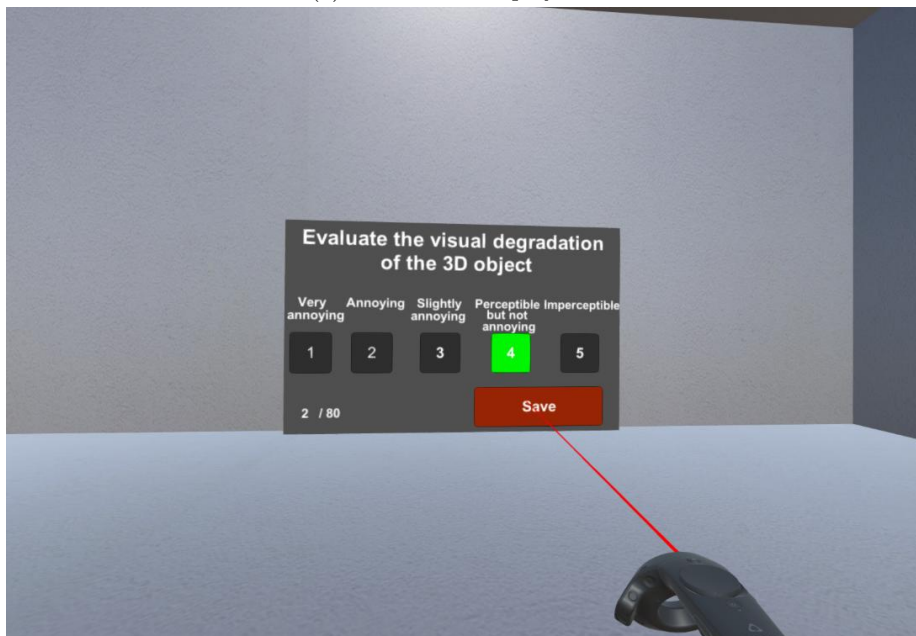


(b) The rating room

Figure 1: The experimental environment of the ACR-HR test



(a) The stimuli display room



(b) The rating room

Figure 2: The experimental environment of the DSIS test

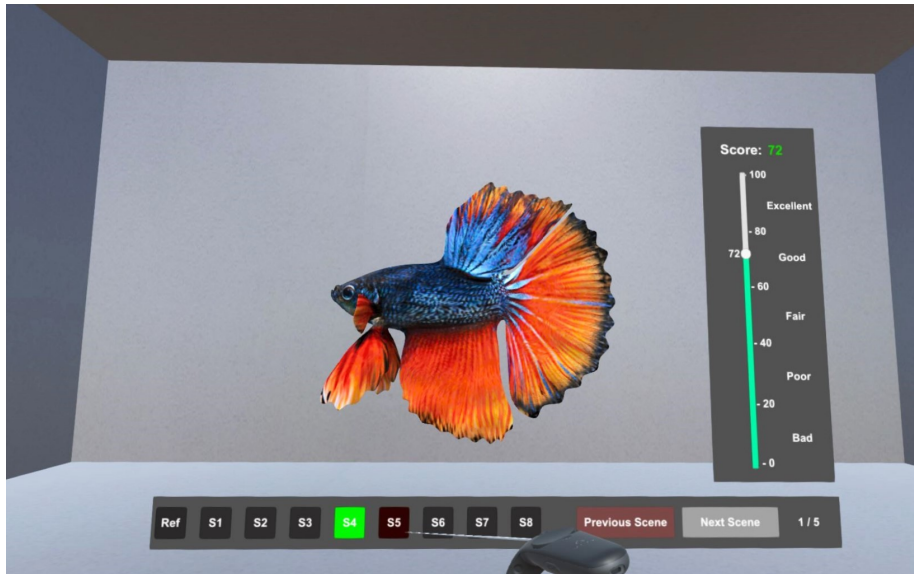


Figure 3: The experimental environment of the SAMVIQ test

## 2 Additional results

In this section, we present, in Figure 4, the computed DMOS of the ACR-HR experiment for all stimulus for both groups G1 and G2. We also present their confidence intervals in Figure 5. Moreover, Figure 6 details the MOS and the CI obtained by the observers of the 2 groups involved in the DSIS tests. Figure 7 summarizes the results of the ACR-HR and DSIS sessions using boxplots of MOS. We recall that G1's subjects did the ACR-HR session first followed by the DSIS session, while G2's subjects did the DSIS session first and then the ACR-HR session. Finally, Figure 8 presents the DMOS and the CI acquired for all the stimuli in the SAMVIQ test.

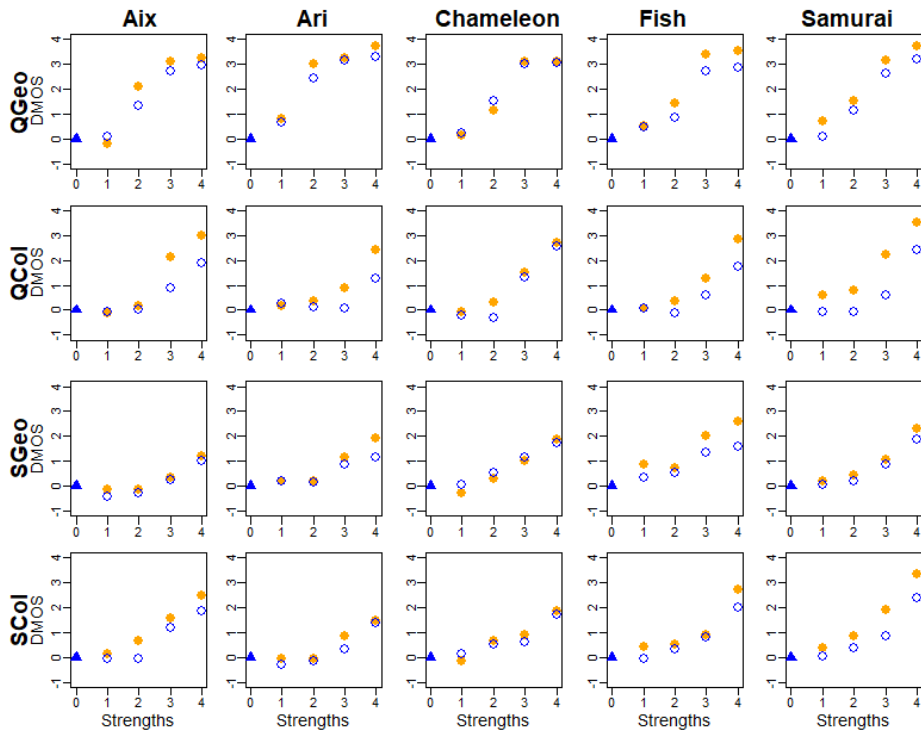
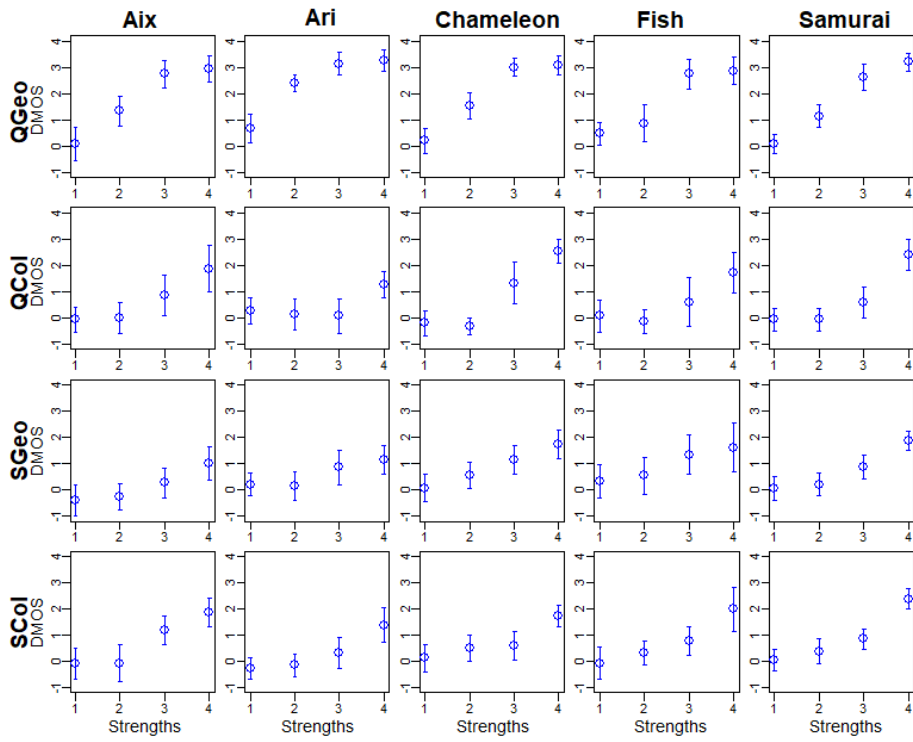
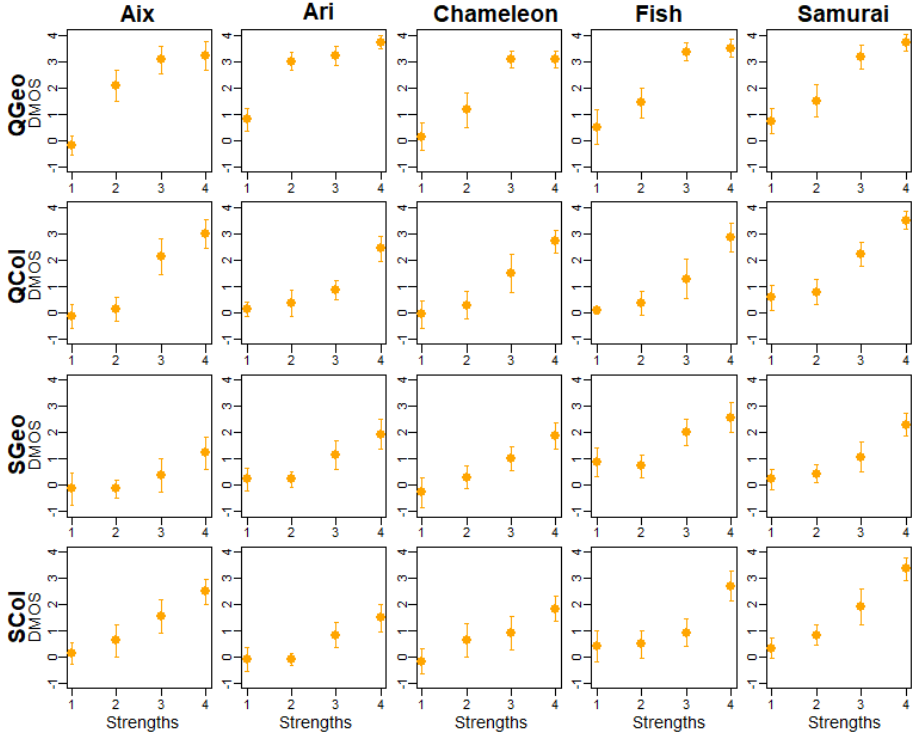


Figure 4: Overview of difference mean scores of the ACR-HR experiment for all stimulus for both groups (the blue and orange dots refer to the DMOS of G1 and G2 respectively).

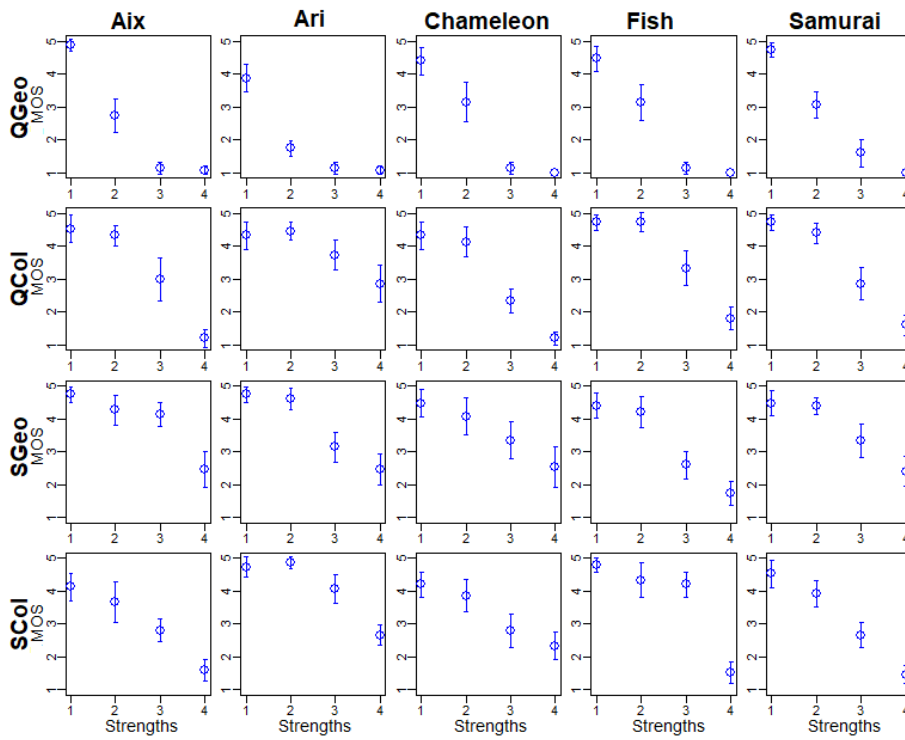


(a) G1

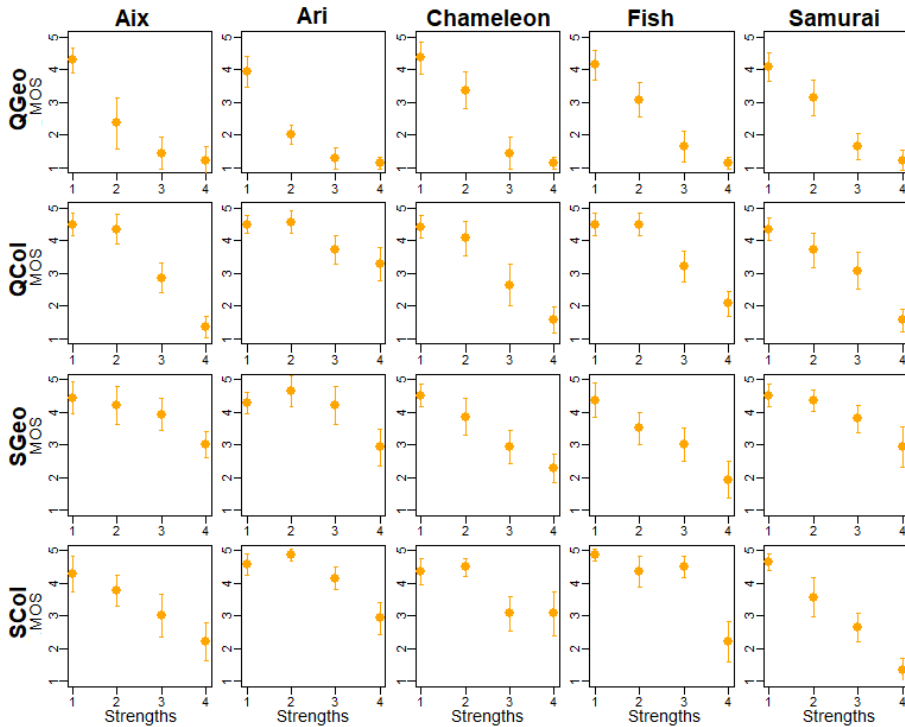


(b) G2

Figure 5: Confidence intervals of the DMOS from the ACR-HR tests for both groups.



(a) G1



(b) G2

Figure 6: Confidence intervals of the MOS from the DSIS tests for both groups.

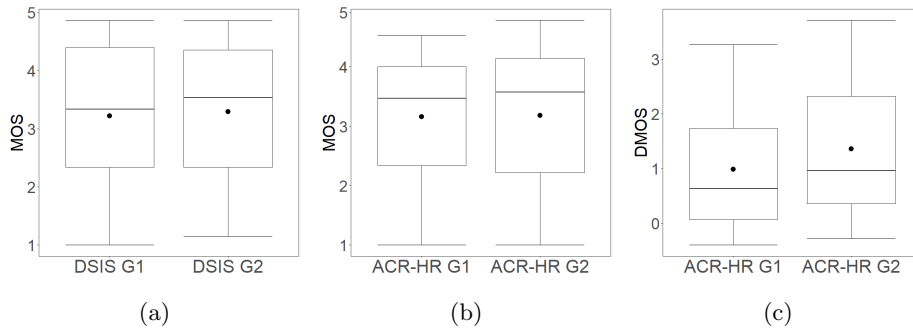


Figure 7: Boxplots of MOSs obtained by the two groups of subjects involved in the ACR-HR and DSIS tests.

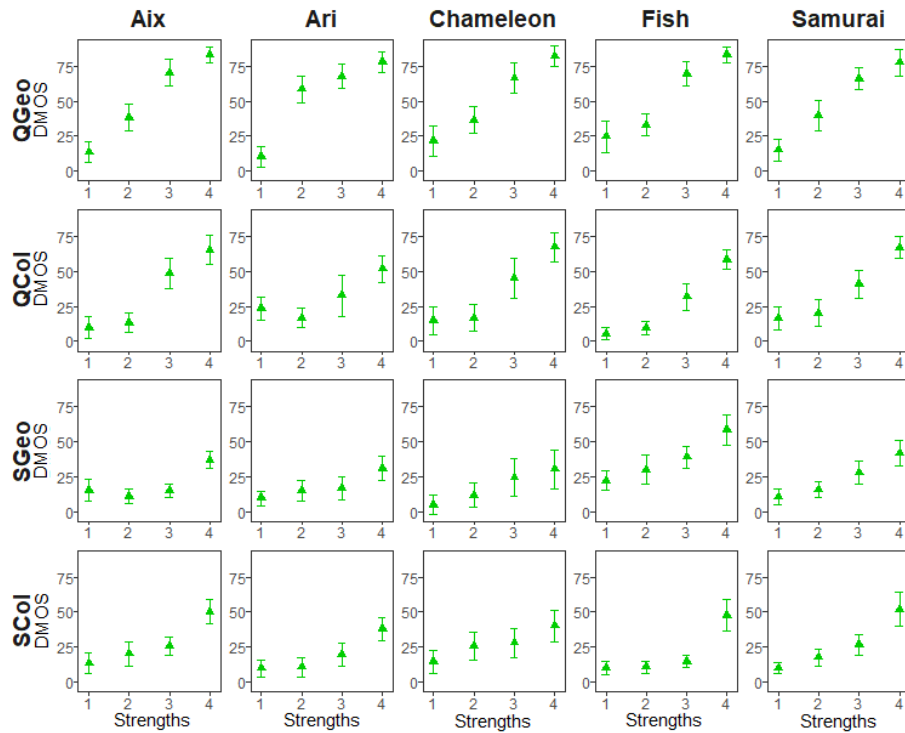


Figure 8: Confidence intervals of the DMOS from the SAMVIQ test

### 3 Intra-rater reliability

In G1 and G2, each participant rated all the stimuli in both the ACR-HR and DSIS tests. Hence, we evaluated, for each participant, the degree of consistency (using ICC) between their scores in the 2 tests. Results are reported in Figure 9, which shows that the consistency between DSIS and ACR scores is higher for G2 observers than for G1 observers. This is coherent with the results observed in section 4.4.3. of the paper (more precisely Figures 5.c and 5.d).



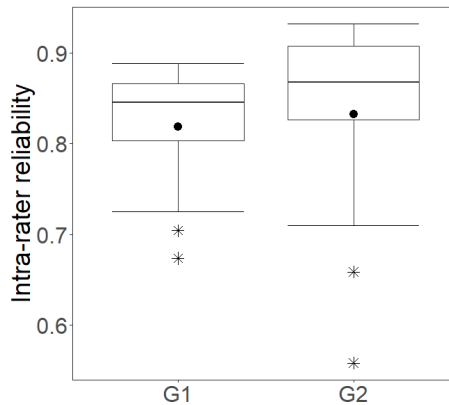


Figure 9: Consistency among the ACR-HR and DSIS tests for the two groups of subjects.

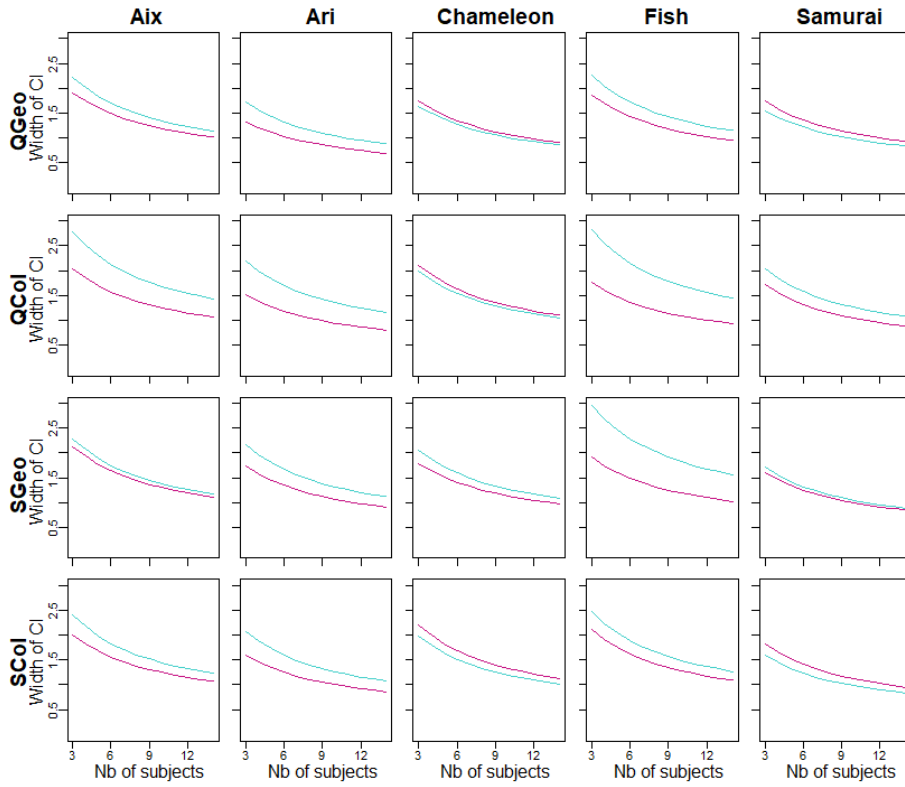
## 4 Confidence intervals

### 4.1 Confidence intervals of ACR-HR and DSIS methods

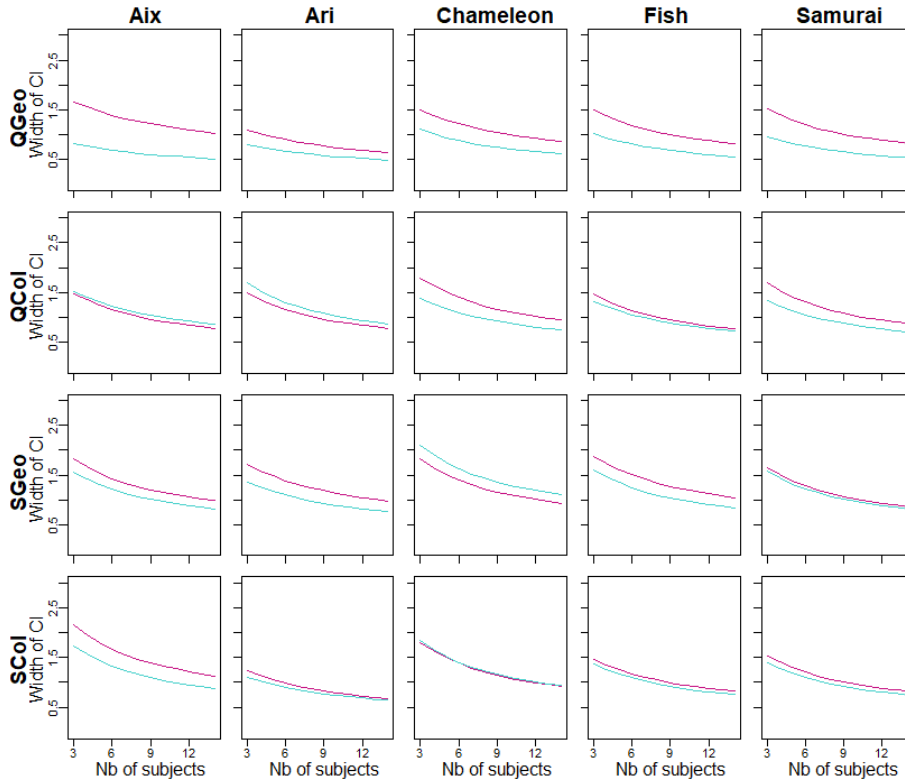
In section 4.4 from the paper, we evaluated the evolution of the width of the confidence intervals (CI) of the ACR-HR and DSIS methods according to the number of subjects. We provide, in Figure 10, a comparison of CI obtained by the observers of the G1 and G2 involved in these 2 tests.

### 4.2 Comparison of the CIs between methods

In this section, we evaluate the performance of the 3 tested methods (ACR-HR, DSIS and SAMVIQ) in terms of the dispersion of individual ratings. Thus we compared the evolution of the CIs according to the number of subjects, among these methods. This comparison was carried out using the ACR-HR scores of G1 and the DSIS scores of G2. We chose these scores because the observers of G1 and G2 first performed the ACR-HR and DSIS sessions respectively and therefore the models were unknown for these subjects, as for the subjects of the SAMVIQ experiment. Results are presented in Figure 11



(a) ACR-HR



(b) DSIS

Figure 10: Mean confidence intervals for both methodologies as a function of the number of observers involved in G1 (turquoise curves) and G2 (violet curves) (G1's subjects did the ACR-HR session 1<sup>st</sup> followed by the DSIS session, while G2's subjects did the DSIS session 1<sup>st</sup> and then the ACR-HR session).

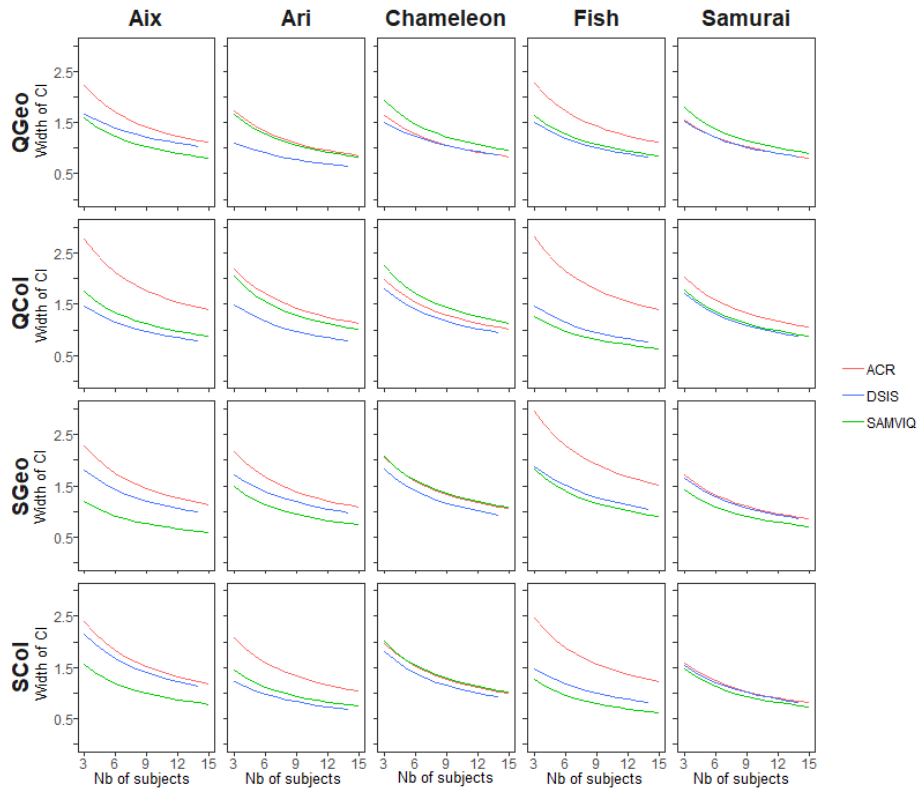


Figure 11: Mean confidence intervals for ACR-HR, DSIS and SAMVIQ methods as a function of the number of observers

## 5 Summary of the tested methods

This section reviews all the experimental details and information of the ACR-HR, DSIS, SAMVIQ methods and pinpoints the main differences between them.

Table 1: Experimental details of ACR-HR, DSIS, SAMVIQ methods.

	ACR-HR	DSIS	SAMVIQ
<b>Explicit reference</b>	No	Yes	Yes
<b>Quality scale</b>	Bad to excellent	Very annoying to imperceptible	Bad to excellent
<b>Scale type</b>	Discrete five-level likert scale	Discrete five-level impairment scale	Continuous quality scale from 0 to 100 (represented by a slider)
<b>Voting</b>	Global quality of the test stimuli, including hidden references	Difference between the test stimulus and the reference simultaneously shown	Global quality of the test stimuli, including explicit references
<b>Presentation of the stimulus</b>	Once	Once	multiple times (random access approach)
<b>Stimulus presentation time</b>	6s	10s	6s
<b>Possibility to change the vote</b>	No	No	Yes
<b>Subjects involved</b>	15	15	17
<b>Display</b>	VR headset*	VR headset*	VR headset*

(\*) 3D meshes were loaded into the VR scene and rotated in real-time.